

Der er tre måder at sætte billeder og tekst sammen på:
2 rigtige og 1 forkert.

Et eyetrack-forsøg førte til en test, der kan afgøre om man ramte rigtigt. Uden udstyr.

Af Niels Heie, Den Grafiske Højskole

I en daglig avis er der omkring 250 artikler, notitser, kommentarer, reportager etc. Under et kan man betegne dem som redaktionelle enheder, der består af mindst en overskrift, men oftest også en manchete, en brødtekst, mellemrubrikker og illustrationer. Halvdelen af de redaktionelle enheder indeholder billeder.

I layoutet disponeres elementerne i størrelshierarkier, der har den virkning at nogle elementer ses før andre. Layoutet er den optiske forudsætning for den rækkefølge læseren opfatter elementerne i. Da billederne næsten altid er store, læses de først, sammen med overskrifterne.

Læsning er en formodningsproces, hvor læseren bevæger sine øjne rundt blandt de redaktionelle enheder og de elementer som de redaktionelle enheder består af. Han eller hun skal finde ud af meningen med de redaktionelle enheder som helhed. Det gør han ved at forstå elementerne et for et, og ved at forstå deres rolle i forhold til hinanden. Læserens job er at forstå meningen med delene som (en) helhed.

Når billede og tekst virker godt sammen forstår læseren deres indholdsmæssige sammenhæng umiddelbart, hvad enten det drejer sig om et billede og en rubrik eller et billede og en billedtekst. Læseren behøver ikke at læse brødteksten først, men forstår indholdet i sammensætningen som et klart opfattet postulat.

Når billede og tekst virker dårligt sammen forstår læseren ikke den indholdsmæssige sammenhæng. Opfattelsen af postulatet er uklar og klarheden opstår først, når læseren har læst brødteksten. Måske havde historien fungeret bedre uden billede.

Da halvdelen af avisens historier indeholder billeder, er halvdelen af historierne som udgangspunkt i fare for at blive opfattet forkert, fordi postulatet ikke opfattes klart. Den uklare opfattelse af postulatet skaber diffuse forventninger til den videre læsning. Forkerte forventninger ødelægger tilslaget – det, der får fisken til at bide på krogen. Som Søren Pagter siger i et interview i SNDS (2006/3 side 7): *"...sammenhængen mellem tekst og billede er vigtig, for hvis nu læseren begynder med en forkert stemning i forhold til historien, så falder det hele til jorden"*.

Man kan læse korrektur på forholdet mellem tekster og billeder

Den store gevinst ved at sætte tekst og billede sammen er, at skriftsprogets abstrakte karakter reduceres, fordi indholdet forankres (1) til en situation og at det billedmæssige udsagn om en eller anden situation bliver bestemt ved hjælp af sproglige kategorier.

Når sammensætningen er god oplever læseren at billedet tager toppen af den sproglige abstraktion og at det sproglige indhold fjerner billedets mangetydighed. Sammensætningen skaber forståelse.

Når sammensætningen er dårlig, oplever læseren ingen sammenhæng, sammensætningen skaber forvirring.

Det problem kan løses ved hjælp af en metode, der gør det muligt at opdage om en sammensætning kan misforstås eller ej. Man kan kalde metoden et korrekturværktøj. Formålet er at kontrollere om indholdet i et postulat, der formidles som en sammensætning af et billede og en tekst, opfattes som helt eller splittet, om man forstår det eller misforstår det.

Teksters og sætningers indbyrdes forhold kan beskrives ved hjælp af begreberne kohærens og kohæsion. Kohærens er betegnelsen for den overordnede sammenhæng i en tekst. Kohæsion er betegnelsen for den underordnede sammenhæng mellem sætninger. Kohærens gør det muligt at forstå alle sætningerne i en tekst som en helhed. Kohæsionen gør det muligt at forstå sammenhængen mellem de enkelte sætninger. Når en tekst er kohærent hænger den sammen som helhed, der opfattes fordi enkeltsætningerne hænger sammen i kraft af kohæsionen.

Kohærens forudsætter kohæsion.

”Korrekturværktøjet” benytter de sproglige begreber kohærens og kohæsion til at beskrive forholdet mellem tekster og billeder. Testen, der skal afgøre om der er indholdsmæssig sammenhæng mellem tekster og billeder er enkel. Den består blot i at man omsætter billedets indhold til en sætning og sætter den sammen med den foreslåede overskrift eller billedtekst. Derved får man to sætninger, der kan give mening i forhold til hinanden eller ej.

Hvis den sætning, der er en oversættelse af billedet, og rubrik eller billedtekst *giver mening sammen* som ledsætninger, *er der også* mening i sammensætningen af billede og overskrift. Når de to tekstelementer er indholdsmæssigt kohærente, er billede og tekst også indholdsmæssigt kohærente.

Hvis sammensætningen af billedsætningen og rubrik eller billedtekst *ikke giver mening sammen* som ledsætninger, *er der ikke* indholdsmæssig kohærens i sammensætningen. Derfor er der ikke indholdsmæssig kohærens i sammensætningen af tekst og billede.

Forudsætningen var et eyetrack-forsøg

Materialet der ligger til grund for metoden fremkom gennem et eyetrack-forsøg, som vi udførte på DGHs eyetrack-laboratorium *Nilpeter Lab*.

Eyetracking er en nyere form for adfærdsregistrering baseret på computerteknologi og neurologisk viden om synets virkningsmåde.

Når man "eyetracker" en person, registrerer man hvordan personen bevæger sine øjne mens vedkommende ser på f.eks. et skærmbillede. De data der indsamles gør det muligt at se, hvilke ruter øjnene gennemløber i forhold til det visuelle stimulus. Man kan videre se med hvilken frekvens og varighed øjnenes *fiksationer* og *saccader* gennemføres.

Fiksationer er perioder, hvor øjnene står relativt stille og "ser" på det samme lille område et stykke tid. Saccader er de perioder, hvor øjnenes fokus flyttes fra et sted til et andet. En typisk fikstion varer mellem 150 – 600 ms. En typisk saccade varer mellem 10 – 150 ms (2). Der er en markant forskel på de to operationer ud over varigheden: under en fikstion er øjnene "seende", under en saccade er de "blinde".

12 + 12 personer så en serie på 24 skærmbilleder, der indeholdt seks forskellig fotografiske motiver vist fire gange. De to grupper så ikke de samme billeder.

Hvert skærmbillede indeholdt et eller tre ens fotos og noget tekst. Teksten bestod af et spørgsmål og tre svarmuligheder. Personerne skulle vælge mellem de tre svarmuligheder og dermed bestemme hvilken sammensætning af tekst og billede, der var mest objektiv, subjektiv, forkert eller følelsesladet. Efter svaret blev et nyt skærmbillede vist og et nyt spørgsmål stillet.

Billederne var udlånt af Dagbladet Politiken, spørgsmål og svar var skrevet af forsøgets designere. (*Forsøget omtales mere detaljeret i 2. del af denne artikel*).

Et typisk skærmbillede, så således ud (fig.1):

Hvilken tekst er mest OBJEKTIV?



- 1) Billede der lukker lyden ude
- 2) Trafikuheld på Janusgade
- 3) Kultransport på hjul

Fig. 1

De tre måder at gøre det på

Forsøget bekræftede, at der er tre måder at sætte tekst og billeder sammen på:

1) *Selvfølgelige sammensætninger*

De kohæsive elementer er umiddelbart tilgængelige. De er direkte observerbare i både billede og tekst. Billedet viser det, som teksten betegner.

I forsøget blev sammensætningen kaldt den objektive.

2) *Usædvanlige sammensætninger*

Nogle af de kohæsive elementer er umiddelbart tilgængelige, men derudover er det nødvendigt for læserens forståelse af sammensætningen, at han selv formoder muligheden af flere. Teksten betegner kun noget af det, som billedet viser.

I forsøget blev sammensætningen kaldt den subjektive.

3) *Meningsløse sammensætninger*

Der er ingen tilgængelige kohæsive elementer, derfor er der heller ingen baggrund for at formode, hvad der kan mangle.

I forsøget blev sammensætningen kaldt den forkerte.

1) Den selvfølgelige sammensætning

Når læseren oplever sammenhængen som objektiv formidler teksten det samme bogstavelige indhold som billedet, forstået sådan at hvis billedet blev oversat til en tekst ville (næsten) alle opleve den som en tautologi, (dvs. udtrykker det samme) i forhold til den tekst billedet blev vist sammen med.

Men selvom billedet og teksten "siger" det samme, er der stadig stor forskel på indholdet i de to. Forskellen er at et billede viser noget, en tekst hævder noget. Billedet viser et udsnit af en konkret situation på et bestemt sted på et bestemt tidspunkt. Den viste situation kunne andre – hvis de havde været tilstede – også have set. Teksten kombinerer nogle ord, der kan forstås som en hævde af muligheden for et eller andet, tænkt eller faktisk trafikuheld, på et eller andet, tænkt eller faktisk tidspunkt, et eller andet, tænkt eller faktisk sted, på en af de veje, der tænkt eller faktisk hedder Janusgade. Hævdelsen kan forstås af andre, hvis de kender sproget.

Billedet viser eksistensen af en bestemt situation, teksten hævder muligheden af en eller andet situation.

Den indholdsmæssige sammenhæng opstår, fordi det giver mest mening at formode, at den ulykke der ses på billedet, er den ulykke der benævnes i teksten.

Kohærensens, dvs. den overordnede sammenhæng, bliver mulig fordi der er elementer i billedet og teksten, der skaber kohæsion, den underordnede sammenhæng. De kohæsive elementer er de betydningselementer, der optræder i både billeder og tekster. I eksemplet med trafikuheldet er der mindst følgende sæt af kohæsive elementer:

- Det man ser kan betegnes med ordet trafikuheld, teksten indeholder det samme ord.
- Stedet hvor det foregår kan betegnes som en gade med et bestemt navn. I teksten nævnes en gade ved navn.
- Situationen på billedet er af samme art, som situationen i teksten.
- Situationen på billedet foregår, på samme måde som situationen i teksten foregår.

Læseren oplever uden problemer indholdsmæssig sammenhæng mellem

et foto der viser et *Trafikuheld på Janusgade*
og sætningen *Trafikuheld på Janusgade*

2) Den usædvanlige sammensætning

Når læseren oplever sammenhængen som subjektiv betyder det, at teksten formidler et indhold som opleves kohærent sammen med den sætning, der beskriver indholdet i billedet bogstaveligt. Men de kohæsive elementer er implicite, *ikke* explicite.

Når læseren kan forstå en mening med de to led i en sætning, der hedder:

Trafikuheld på Janusgade, billede der lukker lyden ude

er det fordi læseren selv forestiller sig de kohæsive elementer, der skal til for at forstå ordene som betydningselementer i den samme overordnede betydningshelhed. Læseren forstår formodentlig sammensætningen af ord således:

(Et foto af et) *Trafikuheld på Janusgade*, (er et) *billede, der lukker lyden ude*.

Når læseren kan opleve fotoets indhold kohærent i forhold til teksten, omsætter hun ubevidst billedets bogstavelige indhold til en ligeså entydig tekst + nogle ord med kohæsiv funktion, der ville gøre sammensætningen umisforståelig (ordene i parentes i eksemplet).

Resultatet er, at læseren ved sin egen meddigtning får en samlet sproglig forestilling, som gør forholdet mellem indholdet i teksten og indholdet i billedet forståeligt, dvs. så billedet og teksten opleves som elementer i den samme betydningsbygning.

Fordi teksten har været kohærent med billedet, har læseren fortolket de indholdsmæssige aspekter af billedet, som om det var en sætning. Denne hypotese om den kognitive proces kan forklare det faktum, at testpersonerne gennemsnitlig brugte 25% længere tid på at løse den subjektive opgave end på den objektive.

3) Den absurde sammensætning

Når læseren oplever at en sammensætning er forkert, er der ingen kohæsive elementer, dvs ingen elementer der findes både i tekster og i billeder, hverken explicite eller implicite.

Læseren oplever ikke nogen mening med de to led i en sætning, der hedder:

Trafikuheld på Janusgade, kultransport på hjul

fordi læseren ikke kan forestille sig, hvilke kohæsive elementer der skulle til for at få indholdet i billede og tekst til at hænge sammen. Derfor oplever han eller hun heller ikke, at der ikke er nogen mening med

(Et foto, der viser et) *Trafikuheld på Janusgade* (og sætningen) *kultransport på hjul*

Læseren står så i den vanskelige situation, at han skal vælge at forstå sammensætningen enten som en fejl, eller som et eksempel på svært tilgængelig dadaisme, eller som et eksempel på en særlig art af kreativ kommunikation, hvor meningen er at læseren skal blive opmærksom på absurditeten.

Idéen som maggi-terning

Forsøgets resultat kan koges ind og omsættes til følgende trefingrede håndfuld af anbefalinger:

A) Tekst og billede hænger indholdsmæssigt sammen, når det giver mening at opfatte

billedets indhold som noget der er

+ **tekstens indhold som noget billedets indhold er**

sammensætningen er objektiv

B) Tekst og billede hænger indholdsmæssigt sammen, når det giver mening at opfatte

billedets indhold som noget der er

+ **tekstens indhold som noget billedets indhold kunne være**

sammensætningen er subjektiv

EKSEMPEL: Hvis et billede forestiller Anders Fogh Rasmussen, Pia Kjærsgård og Bent Bentsen der kommer gående med et stort smil direkte mod fotografen, vil det give mening at sætte billedet sammen med teksten *De tre partiledere fra regeringen og dens støtteparti på vej til møde*. Men det vil også give mening at sætte billedet sammen med teksten *Frisk pust over landet* uanset hvilken situation de tre personer faktisk var i, da billedet blev taget.

I den første version siger billedet (næsten) det samme som teksten: det der er i billedet, er også i teksten. Det giver mening at kalde sammensætningen objektiv.

I den anden skal læseren forestille sig en masse for at finde sammenhængen: f.eks. prøve at forstå indholdet i teksten som noget der kunne være resultatet af de tre personers aktivitet, hvis man var sympatisk indstillet overfor deres politiske holdninger. Det der er i billedet, *kunne være* det der er i teksten. Det giver mening at kalde sammensætningen subjektiv.

C) Tekst og billede hænger ikke sammen, for det giver ikke mening at opfatte en sammenhæng mellem:

billedets indhold som noget der er
+ tekstens indhold som noget billedets indhold *ikke kunne være*

sammensætningen er meningsløs

EKSEMPEL: Et foto viser en stor lastbil væltet omkuld i en grøft. Det har en billedtekst, der hedder: *Valutabalancen er truet*. Indholdet er selvmodsigende. Bilen kunne være væltet, chaufføren kunne have hørt Champagnegaloppen meget højt, bilen kunne have været en skodbil. Men hverken den eller hændelsen eller billedet kunne have været en valutabalance.

Altså giver det mening at kalde sammensætningen meningsløs.

Eyetracking-forsøget

Det eyetrack-forsøg, som korrekturmetoden udsprang af, var det afsluttende projekt på et uddannelsesforløb som DGHs eyetrack-gruppe gennemførte på Lunds Universitet med professor Kenneth Holmquist (3) som vejleder. Formålet var bl.a. at fastlægge tidsforbruget i forbindelse med testpersoners oplevelse af indholdsmæssige forhold mellem tekster og billeder ved hjælp af øjenbevægelsesdata.

Skærbillederne – det testpersonerne så – blev fremstillet i to layouts. Det ene indeholdt et foto med de tre svarmuligheder placeret som linier i en spalte direkte under billedet, fig. 1. Det andet indeholdt tre identiske fotos og de tre svarmuligheder var placeret som billedtekster, et under hvert billede, altså som i et billedtekstlayout, fig. 2. De to layouts blev kaldt henholdsvis spaltelayout (SL) og billedtekstlayout (BL).



Fig 1



Fig 2

Testpersonerne skulle bestemme, hvilken svarmulighed der ”passede” bedst til fotoets indhold, hvis sammensætningen skulle betegnes som henholdsvis mest objektiv, mest subjektiv, mest forkert og mest følelsesladet. Der var ingen tidsbegrænsninger for testpersonen som skulle besvare spørgsmålet, før næste billede blev vist.

Det skærmbillede der indeholdt et billede af et trafikuheld og bl.a. teksten ”*Trafikuheld på Janusgade*” er blevet gennemgået. Et andet eksempel var dette (fig. 2):

Fotografiet viste et ældre bondehus på en øde mark. Testpersonerne kunne vælge mellem disse svarmuligheder:

- *Et lille hus på landet*
- *Fængsel for livstid*
- *32 tomme bakker*

Inden forsøget blev afviklet, havde vi der stod bag forsøget gættet på de forslag testpersonerne ville vælge. F.eks. formodede vi at testpersonerne til billedet af et trafikuheld ville vælge:

- *Trafikuheld på Janusgade* til den mest objektive sammensætning
- *Billede der lukker lyden ude* til den mest subjektive
- *Kultransport på hjul* til den mest forkerte
- *Billede der lukker lyden ude* til den mest følelsesladede

I forhold til billedet af huset på landet formodede vi, at testpersonernes ville vælge teksten

- *Et lille hus på landet* til den mest objektive sammensætning
- *Fængsel på livstid* til den mest subjektive
- *32 tomme bakker* til den mest forkerte
- *Fængsel på livstid* til den mest følelsesladede

Analysen af data

Analysen af eyetrack-dataene foregik efter et system, hvor skærmbillederne blev opdelt i interesseområder (forkortet *AOI*, som betyder *Areas Of Interest*).

Skærmbillederne med spaltelayout indeholdt 5 interesseområder:

- *spørgsmålet*
- *tre svarmuligheder*
- *et foto*

Skærmbillederne med billedtekstlayout indeholdt 7 interesseområder:

- *spørgsmålet*

- tre svarmuligheder
- et foto vist tre gange



Stimulusbilleder med markering af Areas of Interest (de gule felter)

Af de fire spørgsmål analyserede vi kun resultaterne fra de mest distinkte spørgsmål: dem om de objektive og de subjektive sammensætninger. I analysen af forholdet mellem vores forventede og testpersonernes faktiske forslag, inddrog vi også resultaterne fra spørgsmålet om den mest forkerte sammensætning, både den procentvise enighed mellem testpersonerne og os, og tidsforbruget.

Resultatet af analysen blev fremstillet som en samlet gennemsnitsværdi. Værdien fremkom på følgende måde:

Værdierne af det individuelle tidsforbrug i hver AOI, blev adderet og brugt som udgangspunkt for udregningen af en gennemsnitsværdi for hver person, hvert spørgsmål, hvert motiv og hvert layout. I layoutet med de tre ens billeder (BL) havde vi lagt værdierne for de tre fotos sammen i den endelige gennemsnitsværdi. I begge layouts lagde vi værdierne for de tre svarmuligheder sammen i den endelige gennemsnitsværdi.

Resultater

TESTPERSONERNES GENNEMSNITLIGE TIDSFORBRUG

- 18% af tiden på spørgsmål (begge layouts)

En femtedel af tiden blev brugt på at læse og vende tilbage til spørgsmålet i den dynamiske formodningsproces som læseren var i under forsøget.

- 30% af tiden på fotos i layouts med et billede (SL-layout)
- 42% af tiden på fotos i layouts med tre billeder (BL-layout)

Selvom de tre billeder var ens i BL-layoutet brugte læseren væsentlig mere tid på at se på de tre billeder end på det ene (SL-layoutet).

- 44% af tiden på svarene i layouts med et billede (SL-layout)

· 24% af tiden på svarene i layouts med tre billeder (BL-layout)

Hvor der kun var et foto i skærmbilledet brugte testpersonen væsentlig længere tid på svarmulighederne end på svarmulighederne i de skærmbilleder hvor der var tre fotos.

· 10% af tiden blev brugt på tomme områder i layouts med et billede (SL-layout)

· 15% af tiden blev brugt på tomme områder i layouts med tre billeder (BL-layout)

Testpersonerne brugte 10-15% af tiden med at se på noget udenfor for områder med visuelt indhold. Måske har vi her en begyndende målelig beskrivelse af den værdi som "luft" har i et layout.

SAMLET TIDSFORBRUG

TP brugte ca. den samme tid på den objektive sammensætning i de to layouts.

Men tidsforbruget var ca. 10% større, når testpersonerne skulle bestemme den forkerte sammensætning og ca. 25% større, når de skulle bestemme den subjektive sammensætning.

Tallene gælder for begge layouts. Det kan altså måles i tidsforbruget om en kognitiv opgave er kompliceret eller ej. Det skal man selvfølgelig tage højde for, når man tager stilling til f.eks. en artikel og den måde man formulerer den på. Hvor meget skal læseren formode?

SAMMENFALD MELLEML FORVENTEDE OG AFGIVNE SVAR

Sammenfaldet mellem vores valg og testpersonernes valg var relativt høj.

Enighed om 95% af de objektive sammensætninger

Enighed om 78 % af de subjektive sammensætninger

Enighed om 85% af de forkerte sammensætninger

Den høje enighed er vigtig, fordi den bekræfter at det er muligt at formidle visuelt ved hjælp af billeder og tekster sat sammen. Læserne forstår meddelelsen på ca. samme måde som afsenderen.

KONKLUSION

· læseren deler i overvejende grad vores opfattelse af, hvad man kan kalde en objektiv og en subjektiv sammensætning af tekster og billeder.

· det komplekse i den indholdsmæssige relation mellem tekster og billeder har stor betydning for den mængde tid, som læseren bruger på at forstå. Jo sværere opgave, jo mere skal læseren fortolke selv, jo længere tid skal han bruge.

· layoutets udformning påvirker læserens visuelle adfærd. De to layouts gav væsensforskellig visuel læseradfærd.

Læseprocessen er en formodningsproces. Den person, der formidler budskabet skal formulere sig visuelt så læseren misforstår så lidt som muligt. Men formidleren må også indse hvilke budskaber, der skal formuleres så det kræver meget formodningsarbejde at forstå dem og hvilke

der skal kræve lidt. Afsenderen skal gøre sig klart, hvilket fortolkningsarbejde læseren skal præstere for at finde meningen med meddelelsen.

En meddelelse, der er formuleret som en komposition af tekster og billeder, skal kunne læses som en relation mellem tekster og billeder. Ellers bliver den ikke forstået rigtigt.

Når det sker, bider fisken.

(1) Roland Barthes bruger udtrykket *forankring* om forholdet mellem billede og tekst i artiklen *Billedets Retorik* i antologien *Visuel Kommunikation*, Medusa, Kbh. 1964 .

Mellem tekst og billede bruger han udtrykket *afløsning*. Jeg finder begrebet *bestemmelse* bedre.

(2) Andrew Duchowski: *Eye Tracking Methodology, Theory and Practice*

Springer Verlag, 2003, London

(3) Kenneth Holmquist har bl.a. publiceret undersøgelsen *Daily Newspaper Layout – designers prediction of readers' visual behavior – a case study*, Lunds Universitet, 2003

Rapporten om eyetrack-forsøget kan downloades fra DGH's hjemmeside. Titlen er ”*Objektive og subjektive indholdsrelationer mellem sproglige og ikke-sproglige tegn, der bruges i grafiske produkter*”

Teksten findes her:

http://dgh.inforce.dk/graphics/DGH%202006%20tekst/ET_lundrapport_text_0406.pdf

Bilagene findes her:

http://dgh.inforce.dk/graphics/DGH%202006%20tekst/ET_rapport_bilag1_2_3_4NY.pdf

Niels Heie (1945), grafisk designer og underviser på Den Grafiske Højskole siden 1996.

Uddannet som filminstruktør fra Den Danske Filmskole, exam. art. i filmvidenskab fra

Københavns Universitet og grafisk designer fra Den Grafiske Højskole. Har udgivet:

At tale til øjet - en bog om grafisk retorik. Grafisk Litteratur, Kbh. 1999.

Spisesedler – plakater, der sælger aviser. Den Grafiske Højskole, 2004